Amendments to the Specification:

Please replace paragraph 8 with the following rewritten paragraph:

Yet another aspect of the invention is directed to a system comprising a parser component, a context generation component, and a comparator component. The parser component receives search queries and ~~identifues~~ identifies potential stopwords in the search queries. The context generation component generates context data based on the search queries and the potential stopwords. The comparator component compares the context data to determine those of the potential stopwords that effect generation of the context data.

Please replace paragraph 42 with the following rewritten paragraph:

Referring back to Fig. 4, comparator component 405 compares context data corresponding to multiple queries from context generation component 403. Based on the comparison, comparator component 405 determines whether the context data from multiple sets of documents are "substantially similar." Whether a set is substantially similar to another set can then be used by stopword detection component 225 to determine, as described in more detail below, whether to include or exclude the stopword from a final rewritten version [[fo]] of the stopword.

Please replace paragraph 45 with the following rewritten paragraph:

Other techniques, such as those based on the relevance scores returned with each category, could alternatively be used. More specifically, the similarity metric mentioned

in the previous paragraph may be calculated as a weighted metric based on the category relevance scores. For example, the relevance scores associated with each of the ~~categores~~ categories in common between the two sets may be summed and then divided by the sum of all the relevance scores of the different categories in the two sets. Alternatively, the relevance scores between the two sets can be normalized such that the sum for each set, or the sum of squares for each set, is one. The products of the relevance scores of matching categories may then be summed to obtain a similarity metric. A further modification in calculating this similarity metric may be based on additional similarity scores that define similarity between different categories. For example, there may be two categories that are both about slightly different types of cartoons, and the relatedness of these two categories may be defined with a category similarity score. In this situation, the similarity metric may then be calculated based on comparing every pair of categories associated with two queries, computing their similarity scores to each other, ~~multipling~~ multiplying by the relevance scores, adding these values, and then normalizing by dividing by the sum of the relevance scores of the different categories in the two sets.

Please replace paragraph 52 with the following rewritten paragraph:

Returning to the ~~examplary~~ exemplary initial search query "show me the way lyrics," the stopwords identified for this search query may be "show me" and "the" (act 702). Accordingly, S+ would be "show me the way lyrics" and S- could be "* * * way lyrics" (acts 705 and 706). Because S- is a less specific query than S+, it is likely to result in more context data and/or less specific context data. For example, when the

context data includes sets of documents, the documents for S- may refer to songs that contain the term "way" in the title but are not titled "Show Me the Way," such as the songs "My Way" or "Walk this Way." Accordingly, the context data for S- and S+ are likely to be determined to be not substantially similar, (acts 709 and 711), and it would thus be desirable to use S+ as the final search query.